

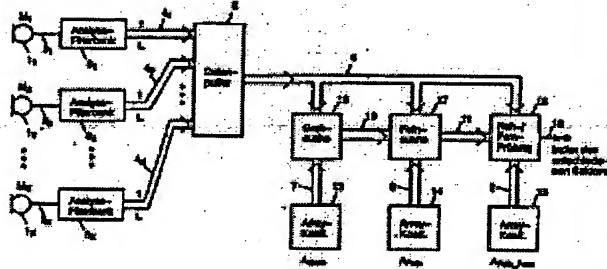
Speaker location method using microphone array

Patent number: DE19540795
Publication date: 1997-05-07
Inventor: ZELINSKI RAINER DR ING (DE)
Applicant: DEUTSCHE TELEKOM AG (DE)
Classification:
- international: H04R1/40; H03H11/36
- european: H04R1/40C; H04R3/00B
Application number: DE19951040795 19951102
Priority number(s): DE19951040795 19951102

[Report a data error here](#)

Abstract of DE19540795

The method involves using the microphone array for the speaker location, such that all the microphone signals of the array are simultaneously evaluated. The speech region is divided into numerous, finite sectors, and the correct sector is selected according to the plausibility criteria. The individual microphone signals, connected for their transit time and phase values, are tested against each other for the phase position coherence, prior to forming the array summation signal. Then the sector with maximum coherence is selected. Preferably during this test, varying frequency band decisions are taken, also according to the maximum coherence.



Data supplied from the **esp@cenet** database - Worldwide

THIS PAGE BLANK (USPTO)

(19) BUNDESREPUBLIK

DEUTSCHLAND



DEUTSCHES
PATENT- UND
MARKENAMT

(12) **Patentschrift**
(10) DE 195 40 795 C2

(51) Int. Cl. 7:
G 01 S 5/20

(21) Aktenzeichen: 195 40 795.4-35
(22) Anmeldetag: 2. 11. 1995
(43) Offenlegungstag: 7. 5. 1997
(45) Veröffentlichungstag
der Patenterteilung: 20. 11. 2003

E4

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

(73) Patentinhaber:

Deutsche Telekom AG, 53113 Bonn, DE

(72) Erfinder:

Zelinski, Rainer, Dr.-Ing., 14193 Berlin, DE

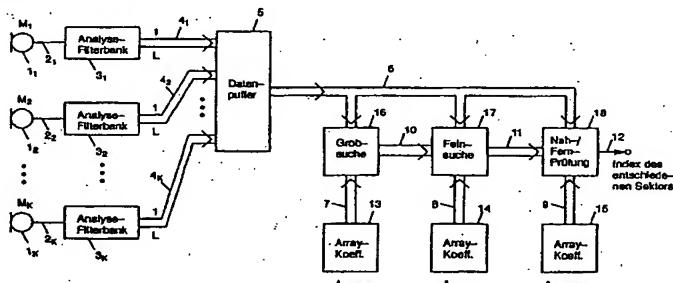
(56) Für die Beurteilung der Patentfähigkeit in Betracht gezogene Druckschriften:

GB 21 15 150 A
GB 20 97 121 A
JP 6-2 61 388 A
JP 1-2 64 487 A

Kellermann, W. "A Self-Steering Digital Microphone Array" In: Int. Comf. Acoust., Speed and Signal Proc. ICASSP '91, Toronto, 1991, S.3581-3584;
Flanagan, J.L. et al.: "Computer-steered microphone arrays for sound transduction in large rooms". In: J. Acoust. Soc. Am. Vol.78, Nov. 1985, S. 1508 -1518;

(54) Verfahren zur Sprecherlokalisierung mit Hilfe eines Mikrofonarrays

(57) Verfahren zur Sprecherlokalisierung mit Hilfe eines Mikrofonarrays, bei dem alle Mikrofonsignale des Arrays gleichzeitig ausgewertet werden und der Sprecherraum in endlich viele Sektoren aufgeteilt und der zutreffende Sektor nach Kriterien der Plausibilität ausgewählt wird, dadurch gekennzeichnet, daß die laufzeit- bzw. phasenkorrigierten, einzelnen Mikrofonsignale vor der Bildung des Array-Summensignals miteinander auf Kohärenz der Phasenlage geprüft werden und danach der Sektor mit maximaler Kohärenz ausgewählt wird.



Beschreibung

[0001] Die Erfindung bezieht sich auf ein Verfahren zur Sprecherlokalisierung mit Hilfe eines Mikrofonarrays der im Oberbegriff des Patentanspruch 1 näher definierten Art. Ein derartiges Verfahren ist z. B. von W. Kellermann: "A selfsteering digital microphone array", Int. Conf. Acoust., Speech and Signal Proc. ICASSP'91, Toronto, 1991, S. 3581–3584 beschrieben.

[0002] Beim Freisprechen in geräuschanbelasteter Umgebung bietet der Einsatz eines Mikrofonarrays eine wirkungsvolle Möglichkeit zur Reduktion des aufgenommenen Geräuschsignals, das dem gewünschten Sprachsignal überlagert ist. Voraussetzung für eine optimale Funktion des Mikrofonarrays ist jedoch, daß die Hauptkeule im Richtdiagramm des Arrays möglichst genau auf die Position des Sprechers ausgerichtet ist.

[0003] Zur Sprecherlokalisierung für den Einsatz von Mikrofonarrays sind eine Reihe von Verfahren in der Literatur beschrieben und zum Teil auch realisiert worden. Die meisten dieser Verfahren beruhen auf einer Kreuzkorrelationsanalyse von jeweils einem Paar von Mikrofonsignalen. Aus der zeitlichen Lage des Maximums der Kreuzkorrelation kann auf den Zeitversatz des Sprachsignals im Signalpaar geschlossen werden. Eine Fortführung der Kreuzkorrelationsanalyse für andere Signalpaare mit anschließender Auswertung ermöglicht dann einen Rückschluß auf die mutmaßliche Sprecherposition. Der erforderliche Rechenaufwand ist nicht sehr hoch, nachteilig ist jedoch die erhöhte Fehleranfälligkeit bei ungünstigen SNR-Werten.

[0004] Eine wesentlich robustere Schätzung der Sprecherposition ist möglich, wenn nicht einzelne Signalpaare, sondern die Kombination aller Mikrofonsignale aus dem Mikrofonarray gleichzeitig ausgewertet wird. Dazu gibt es einen Vorschlag in der Dissertation "Geräuschreduktion bei Sprachübertragung mit Hilfe von Mikrofonarraysystemen"/S. Gierl, Universität Karlsruhe, 1990, der dort als "Fehlerfunktionsmethode" beschrieben wird. Von großem Nachteil ist jedoch der dazu erforderliche, extrem hohe Rechenaufwand, der eine Echtzeitrealisierung mit heute verfügbarer Technologie nicht gestattet.

[0005] Ein anderes Konzept zur Schätzung der Sprecherposition, das ebenfalls auf einer gleichzeitigen Auswertung aller Mikrofonkanäle beruht, ist bereits mit heutiger Technologie in Echtzeit realisierbar. Hierbei wird der dreidimensionale Raum, der die zulässigen Sprecherpositionen beschreibt, in endlich viele Sektoren unterteilt. Für jeden Sektor ist ein Satz von Parameterwerten zur Richtungssteuerung des Mikrofonarrays abgespeichert, der dem räumlichen Zentrum des jeweiligen Sektors zugeordnet ist. Zur Durchführung der Sprecherlokalisierung werden alle Sektoren abgeprüft, und das jeweilige Array-Ausgangssignal wird analysiert. Entschieden wird für denjenigen Sektor, für den sich die größte Plausibilität aufgrund eines ausgewählten Kriteriums ergibt. Hierzu gibt es verschiedene Vorschläge, z. B. die Suche nach der maximalen Signalamplitude unter Einhaltung weiterer Randbedingungen (J. L. Flanagan et al.: "Computersteered microphone arrays for sound transduction in large rooms". J. Acoust. Soc. Am., Vol. 78, Nov. 1985, S. 1508–1518) oder gar der Einsatz von Verfahren der Mustererkennung (W. Kellermann: "A selfsteering digital microphone array". Int. Conf. Acoust., Speech and Signal Proc. ICASSP'91, Toronto, 1991, S. 3581–3584). Diese Auswerteverfahren sind jedoch primär für Fernfeldbedingungen optimiert, also bei größerem Abstand zwischen Sprecher und Mikrofonarray. Weiterhin sind gewisse Kenntnisse über die Statistik von Sprach- und Störsignal erforderlich, um zu vermeiden, daß eine Störquelle als vermeintli-

ches Sprachsignal detektiert wird.

[0006] Aufgabe der Erfindung ist es, ein Verfahren zu schaffen, das eine Sprecherlokalisierung im Nahfeld, auch bei starken Störschallquellen in der Umgebung des Sprechers, mit geringem Zeit- und Rechenaufwand ermöglicht.

[0007] Die Erfindung löst diese Aufgabe mit dem im Patentanspruch 1 aufgeführten Verfahren.

[0008] Mit den in den Unteransprüchen 2 bis 5 angeführten, vorteilhaften Weiterbildungen des Verfahrens wird die Sicherheit gegenüber Fehleinstellungen und Störbeeinflus- sungen weiter erhöht.

[0009] Die Erfindung wird nachfolgend anhand von Ausführungsbeispielen näher erläutert. In den zugehörigen Zeichnungen zeigen die:

[0010] Fig. 1 ein vereinfachtes Gesamtblockbild,

[0011] Fig. 2 das Blockbild zur Kohärenzmaß-Bestim- mung,

[0012] Fig. 3 die Details zur Grobsuche,

[0013] Fig. 4 die Details zur Feinsuche und

[0014] Fig. 5 die Details zur Nah-/Fern-Prüfung.

[0015] Im Gegensatz zu den oben beschriebenen Verfahren unter Annahme von Fernfeldbedingungen ist das in dieser Erfindung beschriebene Verfahren zur Sprecherlokalisierung für das Freisprechen im Nahfeld optimiert, also für Ab-

stände zwischen Sprecher und Mikrofonarray etwa im Bereich von 0,3 . . . 0,8 m. Basis ist zunächst wieder die Unterteilung des Sprecherraumes in endlich viele Sektoren. Die Auswertung der Array-Ausgangssignale erfolgt jedoch nach einem neuen, bisher noch nicht vorgeschlagenen Verfahren.

[0016] Beim Freisprechen im Nahfeld kann ausgenutzt werden, daß der Direktschall des Sprachsignals vom gewünschten Sprecher dominiert gegenüber Sprach-Schallanteilen (und auch Störschallanteilen) infolge von Reflexionen an den Raumbegrenzungsfächern. Wird also bei der Sprecherlokalisierung gerade der "richtige" Raumsektor abgeprüft, so weisen die laufzeit- bzw. phasenkorrigierten Mikrofonsignale unmittelbar vor Bildung des Array-Summensignals bezüglich des Sprachsignals alle die gleiche Phasenlage auf ("kohärenter Schalleinfall"). Bei der Auswertung der Arraysignale wird nun ein Kohärenzmaß bestimmt, das um so größer ist, je geringer die Phasenunterschiede der korrigierten Mikrofonsignale sind. Bei der Sprecherlokalisierung wird dann für denjenigen Sektor entschieden, bei dem sich das größte Kohärenzmaß einstellt.

[0017] Der wesentliche Vorteil des hier beschriebenen Verfahrens besteht darin, daß eine Kenntnis über die Statistik von Sprach- und Störsignal nicht erforderlich ist. Das Verfahren arbeitet wirkungsvoll für beliebige Sprecher und auch für beliebige, störende Geräuschquellen in der Umgebung des Sprechers. Die Geräuschquellen dürfen dabei auch aus Sprachsignalen bestehen. Das gewünschte Sprachsignal bzw. die dazugehörige Sprecherposition wird allein auf Grund der Kohärenzeigenschaften im Nahfeld bestimmt.

[0018] Bei dem in der Erfindung beschriebenen Verfahren erfolgt die Analyse der Signale des Mikrofonarrays im Frequenzbereich. Durch eine automatische Steuerung wird gewährleistet, daß nur solche Frequenzbänder zur Bestimmung des Kohärenzmaßes berücksichtigt werden, die am wenigsten gestört sind, also die größten Kohärenzbeiträge beisteuern. Der Vorteil dieser Vorgehensweise besteht darin, daß eine sichere Sprecherlokalisierung auch noch bei sehr starken Umgebungsgeräuschen möglich ist.

[0019] Zu der vorliegenden Erfindung wird eine Schaltungsanordnung beschrieben, die eine ausreichend genaue Lokalisierung des Sprechers auch bei sehr ungünstigen Signal-zu-Rausch-Abständen (SNR) ermöglicht, ohne Kenntnisse über die Statistik von Sprach- bzw. Geräuschsignal vorauszusetzen.

[0020] Im vereinfachten Gesamtblockbild (Fig. 1) der Schaltungsanordnung besteht das Mikrofonarray aus K Mikrofonen $1_1 \dots 1_K$. Jedes der Mikrofonsignale $2_1 \dots 2_K$ wird einer eigenen Analyse-Filterbank zugeführt. Die Analyse-Filterbänke $3_1 \dots 3_K$ führen jeweils eine Transformation des Zeitsignals in den Frequenzbereich durch und liefern K Frequenzbereichsignale $4_1 \dots 4_K$. Die Analyse-Filterbank realisiert eine Frequenzbandzerlegung des Gesamtbandes in L Teilbänder. Dann besteht jedes der Signale $4_1 \dots 4_K$ aus L Komponenten, die den komplexen Fourier-Koeffizienten $F_1 \dots F_L$ bei einer Spektralanalyse entsprechen. Die Berechnung der Signale $4_1 \dots 4_K$ wird zeitperiodisch wiederholt, und die Signalwerte über N Zeitperioden werden im Datenpuffer 5 zwischengespeichert. Der Datenpuffer 5 enthält also die komplexen Frequenzbereichs-Signalwerte geordnet über N Zeitpunkte, L Frequenzbänder und K Mikrofonkanäle.

[0021] Zur Reduktion des Gesamt-Rechenaufwandes wird die Suche nach dem zu bestimmenden Sektor, in dem der Sprecher sich am wahrscheinlichsten aufhält, zweistufig durchgeführt: Zunächst wird nur die ungefähre Position ermittelt, wobei ein grobes Sektorraster mit relativ wenigen Sektoren (Anzahl SG) zugrunde gelegt wird ("Grobsuche"). Anschließend wird ein wesentlich feiner abgestuftes Sektorraster um die in der ersten Stufe ermittelte Position herumgelegt mit SF Feinsektoren ("Feinsuche").

[0022] Die Grobsuche erfolgt mittels der Schaltungsanordnung 16, in welche die komplexen Fourier-Koeffizienten 6 eingespeist werden. Im Parameterspeicher 13 befinden sich die komplexen Array-Koeffizienten A_{Grob} , die den möglichen Sprecherpositionen im Grob-Sektorraster zugeordnet sind. Für jeden der SG Grobsektoren und jedes der L Teilbänder ist ein eigener Satz von komplexen Array-Koeffizienten $A_1 \dots A_K$ festgelegt, der jeweils über die Leitung 7 in die Stufe 16 eingespeist wird. Die Phasenwinkel der Array-Koeffizienten sind so berechnet, daß nach Durchführung des Laufzeitausgleichs in der Stufe 16 die Signalphasen in den Kanälen 1 ... K pro Teilband untereinander gleich sind, wenn der Sprecher sich tatsächlich in der zugrunde gelegten Sektor-Position aufhalten würde.

[0023] Die getroffene Vorentscheidung über einen Sektor im Grobraster wird über den Eingang 10 der Schaltungsanordnung 17 (Feinsuche) zugeführt. Im Parameterspeicher 14 befinden sich die komplexen Array-Koeffizienten A_{Fein} , die den möglichen Sprecherpositionen im Fein-Sektorraster zugeordnet sind. Jedem der SG Grobsektoren ist ein Satz von SF Feinsektoren zugeordnet (die Feinsektoren bilden eine Unterteilung des betreffenden Grobsektors), wobei zu jedem Feinsektor die komplexen Array-Koeffizienten für K Kanäle und L Teilbänder gespeichert sind. Aus den Array-Koeffizienten 8, die nach dem gleichen Kriterium wie bei der Grobsuche berechnet wurden, und den Fourier-Koeffizienten 6 wird in der Stufe 17 der optimale Sektor im Feinraster – also die mutmaßliche Sprecherposition – ermittelt. Die getroffene Entscheidung 11 wird der Schaltungsanordnung 18 (Nah-/Fern-Prüfung) zugeführt.

[0024] In der Stufe "Nah-/Fern-Prüfung" wird überprüft, ob die geortete Schallquelle sich tatsächlich im Nahfeldbereich befindet. Dadurch kann verhindert werden, daß das Mikrofonarray auf eine energiereiche Störquelle (z. B. ein lauter Sprecher im Hintergrund) ausgerichtet wird, die jedoch – aufgrund ihres größeren Abstandes zum Mikrofonarray – von vornherein als zulässige Sprachquelle ausgeschlossen werden kann. Die Array-Koeffizienten A_{Fein_Fern} im Datenspeicher 15 sind für Fernfeldbedingungen berechnet. Jedem Sektor im Feinraster ist hierbei ein Sektor mit den gleichen Einfallswinkeln zum Mikrofonarray zugeordnet, jedoch mit dem Abstand zwischen Schallquelle und Mi-

krofonarray gegen Unendlich. Die Fourier-Koeffizienten 6 und die Array-Koeffizienten 9 werden der Stufe 18 zugeführt. Hier wird überprüft, ob für den aus der Stufe 17 bestimmten Feinraster-Sektor eine Nahfeld- oder Fernfeld-Situation vorliegt. Ist eine Fernfeld-Situation wahrscheinlicher, so wird die aus der Feinsuche getroffene Entscheidung 11 nicht akzeptiert. Ist jedoch die Nahfeld-Situation wahrscheinlicher, so wird als Ergebnis 12 der Sprecherortung der Index des entschiedenen Sektors im Feinrastermaß ausgegeben.

[0025] Die Sektorentscheidung in den Stufen 16, 17 und 18 beruht jeweils auf der Bestimmung eines Kohärenzmaßes. Entschieden wird für denjenigen Sektor, für den das Kohärenzmaß am größten ist.

[0026] Die in Fig. 2 dargestellte Schaltungsanordnung zur Kohärenzmaß-Bestimmung (Stufe KOMB) führt eine Auswertung pro Teilband und pro Sektor durch und kommt in den Stufen 16, 17 und 18 zur Anwendung.

[0027] Die komplexen Fourier-Koeffizienten 26 ($F_1 \dots F_K$) gelangen über einen Datenpuffer 21 und einen Schalter 23 zur Betragsbildung in die Stufe 29 mit anschließender Summation über K Kanäle 31 und darauffolgender Quadrierung 33. Das Ergebnis 34 wird in einem Datenspeicher 35 zwischengespeichert und kann bei Bedarf als Signalwert 36 ausgegeben und über den Schalter 37 weiterverarbeitet werden.

[0028] Diese Vorgehensweise erspart Rechenaufwand, da die Fourier-Koeffizienten 26 zwar vom Teilband, nicht aber vom Sektor abhängen, also nur einmal mit Gültigkeit für alle Sektoren ausgewertet werden.

[0029] Die komplexen Array-Koeffizienten 27 ($A_1 \dots A_K$) gelangen über einen Datenpuffer 22 und einen Schalter 24 als Signal 40 in die Multiplizierstufe 41; in der pro Kanal eine Multiplikation mit den entsprechenden Fourier-Koeffizienten 28 erfolgt. Das Ergebnis 42 wird in der Stufe 43 über alle Kanäle summiert und von dem Ausgangssignal 44 wird in 45 das Betragsquadrat gebildet. Das Ergebnis 46 wird zusammen mit dem Signal 38 aus der Verarbeitung der Fourier-Koeffizienten der Divisionsstufe 39 zugeführt. Das Ergebnis 47 ist der Kohärenzwert:

$$KOW = \left| \sum_{i=1}^K A_i F_i \right|^2 / \left[\sum_{i=1}^K |F_i|^2 \right],$$

der einen Wert zwischen 0 und 1 annimmt, sofern die Array-Koeffizienten $A_1 \dots A_K$ auf den Betrag Eins normiert sind. Je geringer der Abstand zwischen dem geprüften Sektor und der tatsächlichen Sprecherposition ist, um so größer wird der Kohärenzwert KOW. In der Stufe 48 erfolgt eine zeitliche Mitteilung von KOW durch Summation über N aufeinanderfolgende Zeitpunkte. Der gemittelte Wert 49 wird anschließend mit dem vom Teilband abhängigen Gewichtsfaktor 50 gewichtet. Diese Gewichtung ermöglicht es, weniger zuverlässige Kohärenzwerte abzuschwächen, z. B. solche für tiefe Frequenzen, bei denen der Abstand zwischen zwei Mikrofonen erheblich kleiner als die Wellenlänge ist. Als Ergebnis wird das Signal 51 ausgegeben, das als Kohärenzmaß KOM bezeichnet wird und das pro Teilband und pro Sektor ermittelt wird.

[0030] Fig. 3 zeigt die Details zur Grobsuche in der Stufe 16. Aus den Fourier-Koeffizienten 6 und den Array-Koeffizienten 7 wird in der Stufe 60 (KOMB nach Fig. 2) das Kohärenzmaß KOM berechnet und als Signal 62 ausgegeben.

[0031] Das Maß KOM wird für alle L Teilbänder jeweils pro zu prüfenden Sektor ausgegeben; die Reihenfolge zur Bearbeitung der Koeffizienten wird in der Stufe 61 über die Leitungen 63₁ und 63₂ und die Schalter 64 und 71 gesteuert. In der

Stufe 66 werden die M größten Werte von KOM aus den Kohärenzmaßen $65_1 \dots 65_M$ für den zu prüfenden Sektor bestimmt. Die Indices 67 der dazugehörigen M Teilbänder werden im Datenspeicher 670 zwischengespeichert.

[0031] In der Stufe 69 werden die ermittelten M größten Werte von KOM, $68_1 \dots 68_M$ summiert, und der Summenwert 70 wird über den Schalter 71 der Stufe 73 zugeführt. Die so beschriebene Signalauswertung wird entsprechend für alle SG Grobsektoren ausgeführt, so daß an der Stufe 73 die Eingangssignale $72_1 \dots 72_{SG}$ anliegen. In 73 wird der größte aus den SG Werten bestimmt und als Signal 77 der Vergleichsschaltung 79 zugeführt. Der dazugehörige Index des Grobsektors 74 gelangt zum Schalter 75 und zum Datenspeicher 670. Ist das Kohärenzmaß \overline{KOM}_G auf Leitung 77 größer als ein vorzugebender Schwellwert 78, so wird das Ergebnis der Sektorschreibe im Grobraster akzeptiert und mit dem Schalter 75 als Signal 76 (Index des entschiedenen Grobsektors) ausgegeben. Gleichzeitig werden aus dem Datenspeicher 670 diejenigen Indices der M Teilbänder ausgewählt und als Signal 671 ausgegeben, die dem entschiedenen Grobsektor 76 über die Verbindung 67 zugeordnet wurden.

[0032] Fig. 4 zeigt die Details zur Feinsuche in der Stufe 17. Aus den Fourier-Koeffizienten 6 und den Array-Koeffizienten 8 wird in der Stufe 80 (KOMB nach Fig. 2) das Kohärenzmaß KOM berechnet und als Signal 82 ausgegeben. Das Maß KOM wird nur für die durch 671 festgelegten M Teilbänder berechnet.

[0033] Diese Berechnung erfolgt für alle zu prüfenden Feinsektoren.

[0034] Durch Einspeisen von 671 und 76 in die Stufe 81 wird die Auswahl der dazugehörigen Teilmenge der Daten gesteuert. Die Reihenfolge zur Bearbeitung der Koeffizienten wird über die Ausgangssignale 83_1 und 83_2 und die Schalter 84 und 88 gesteuert. Die Summation der M Kohärenzmaße $85_1 \dots 85_M$ in der Stufe 86 führt auf den Summenwert 87, der über den Schalter 88 der Stufe 90 zugeführt wird. Diese Signalauswertung wird entsprechend für alle SF Feinsektoren ausgeführt, so daß an der Stufe 90 die Eingangssignale $89_1 \dots 89_{SF}$ anliegen. In 90 wird der größte aus den SF Werten bestimmt und als Kohärenzmaß 92 (KOM_F) ausgegeben. Der dazugehörige Index dieses Feinsektors erscheint als Signal 91 am Ausgang der Stufe zur Feinsuche.

[0035] Fig. 5 zeigt die Details zur Nah-/Fern-Prüfung in der Stufe 18. Aus den Fourier-Koeffizienten 6 und den Array-Koeffizienten 9 wird in der Stufe 100 (KOMB nach Fig. 2) das Kohärenzmaß KOM berechnet und als Signal 102 ausgegeben. Das Maß KOM wird nur für die durch 671 festgelegten M Teilbänder berechnet; diese Berechnung erfolgt nur für den Sektor, der durch den Index 91 festgelegt ist. Durch Einspeisen von 671 und 91 in die Stufe 101 wird die Auswahl der dazugehörigen Teilmenge der Daten gesteuert und der über 103 kontrollierte Schalter 104 betätigt. Die Summation der M Kohärenzmaße $105_1 \dots 105_M$ und der Stufe 106 führt auf den Summenwert 107, der das Kohärenzmaß \overline{KOM}_{FF} für das Feinraster-Fernfeld repräsentiert. Die Signale 107 und 92 (letzterer repräsentiert das Kohärenzmaß \overline{KOM}_F für das Feinraster-Nahfeld) werden in die Vergleichsschaltung 108 eingespeist. Ist \overline{KOM}_F größer als \overline{KOM}_{FF} , so wird die Feinsektor-Entscheidung akzeptiert und über den Schalter 110 als Ergebnis 12 (Index des entschiedenen Feinsektors) angegeben.

raum in endlich viele Sektoren aufgeteilt und der zutreffende Sektor nach Kriterien der Plausibilität ausgewählt wird, dadurch gekennzeichnet, daß die laufzeit- bzw. phasenkorrigierten, einzelnen Mikrofonsignale vor der Bildung des Array-Summensignals miteinander auf Kohärenz der Phasenlage geprüft werden und danach der Sektor mit maximaler Kohärenz ausgewählt wird.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß während der Prüfung auf Kohärenz der Phasenlage variierende Frequenzbandbeschränkungen vorgenommen werden, wobei die Frequenzbandbeschränkung ebenfalls nach dem Kriterium der maximalen Kohärenz durchgeführt wird.

3. Verfahren nach Anspruch 1 und 2, dadurch gekennzeichnet, daß die Prüfung auf Kohärenz der Phasenlage in eine Grobsuche mit geringer Unterteilung in wenige Sektoren und eine anschließende Feinsuche mit stärkerer Unterteilung der bei der Grobsuche ermittelten Sektoren mit zulässigen Sprecherpositionen im Nahbereich des Arrays aufgeteilt wird.

4. Verfahren nach Anspruch 1 bis 3, dadurch gekennzeichnet, daß für die Prüfung der Zulässigkeit des Sektors der Sprecherposition ein zusätzlicher Kohärenzvergleich mit einer Schallquellenposition durchgeführt wird, die den gleichen Raumwinkel zum Mikrofonarray gemäß dem Ergebnis der Grob- und Feinsuche aufweist, jedoch dem Fernbereich des Arrays zugehört.

5. Verfahren nach Anspruch 1 bis 4, dadurch gekennzeichnet, daß während der Prüfung auf Kohärenz der Phasenlage eine Vielzahl von Daten komplexer Frequenzbereichs-Signalwerte gebildet, gespeichert und mit ebenfalls gespeicherten, komplexen Koeffizienten zur Phasenkorrektur bewertet werden, die der jeweiligen zu prüfenden Sprecherposition zugeordnet sind.

Hierzu 5 Seite(n) Zeichnungen

Patentansprüche

1. Verfahren zur Sprecherlokalisierung mit Hilfe eines Mikrofonarrays, bei dem alle Mikrofonsignale des Arrays gleichzeitig ausgewertet werden und der Sprecher-

- Leerseite -

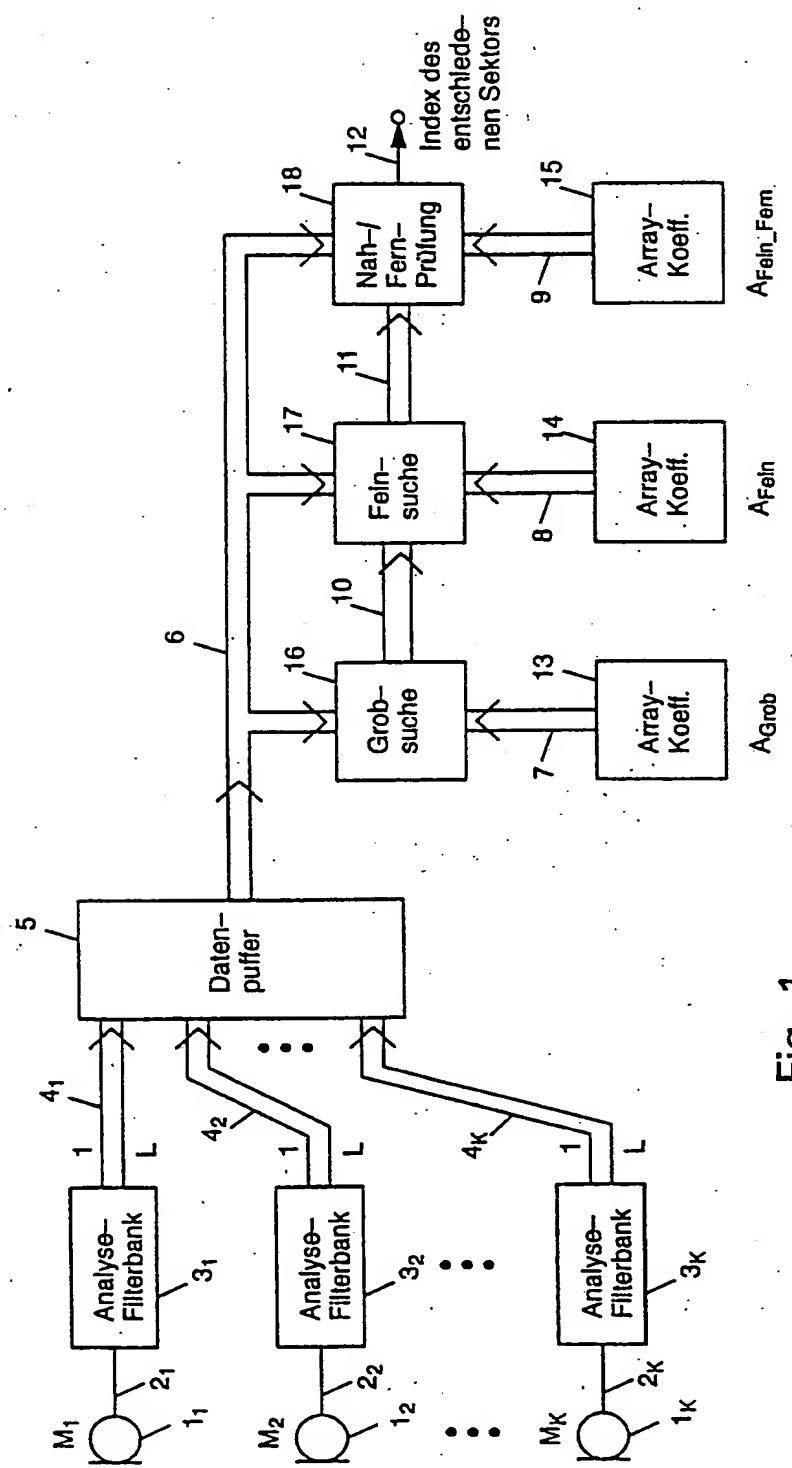


Fig. 1

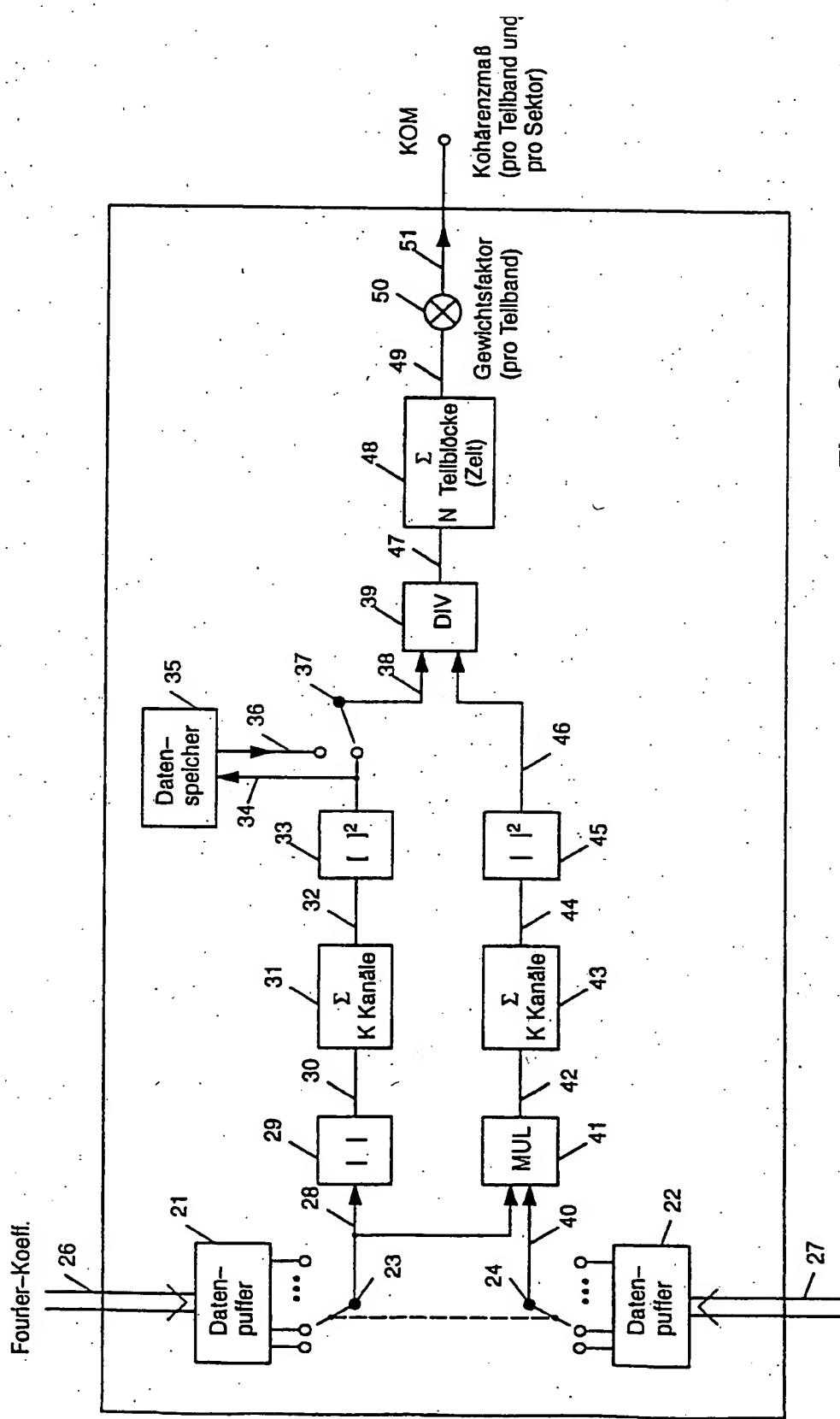


Fig. 2

Array-Koeff.

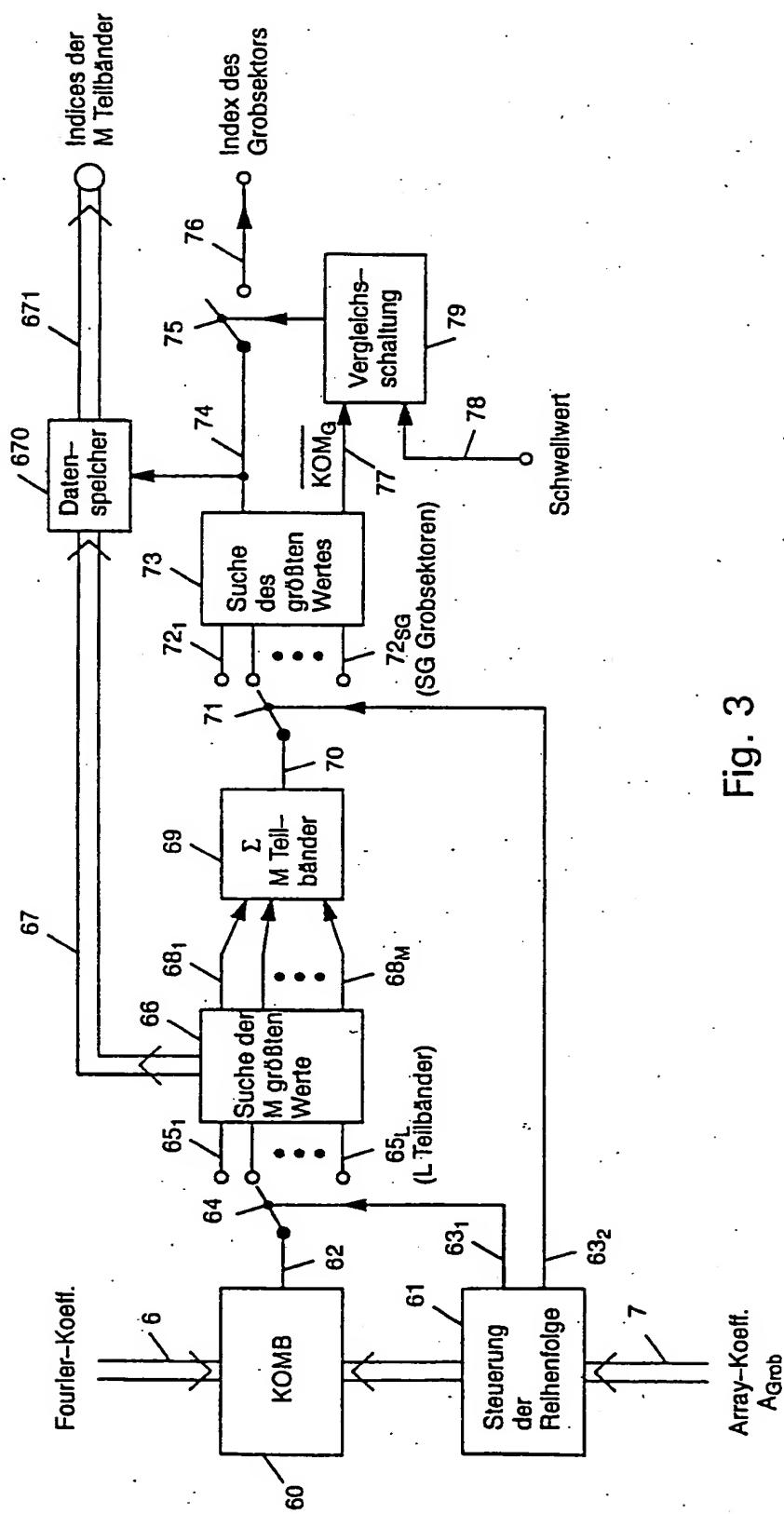


Fig. 3

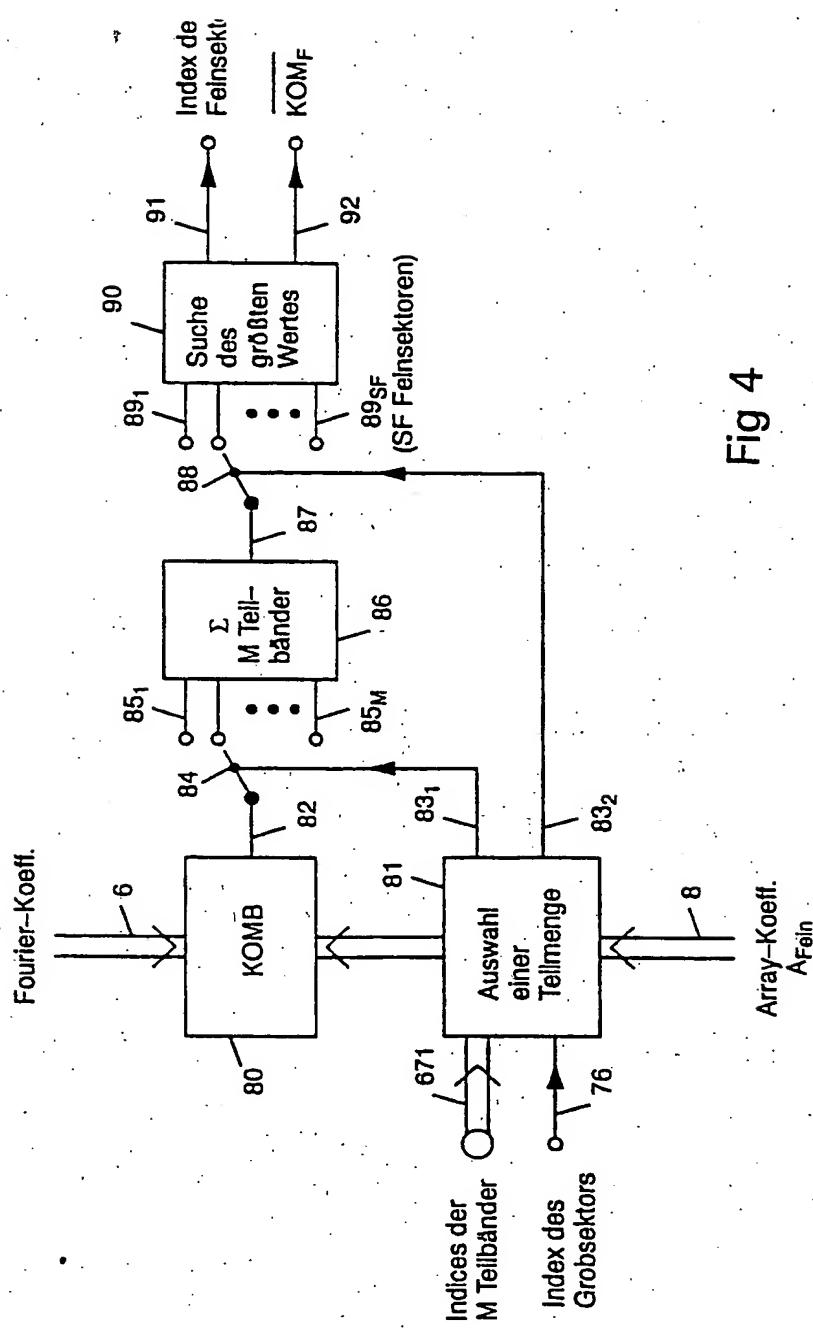


Fig 4

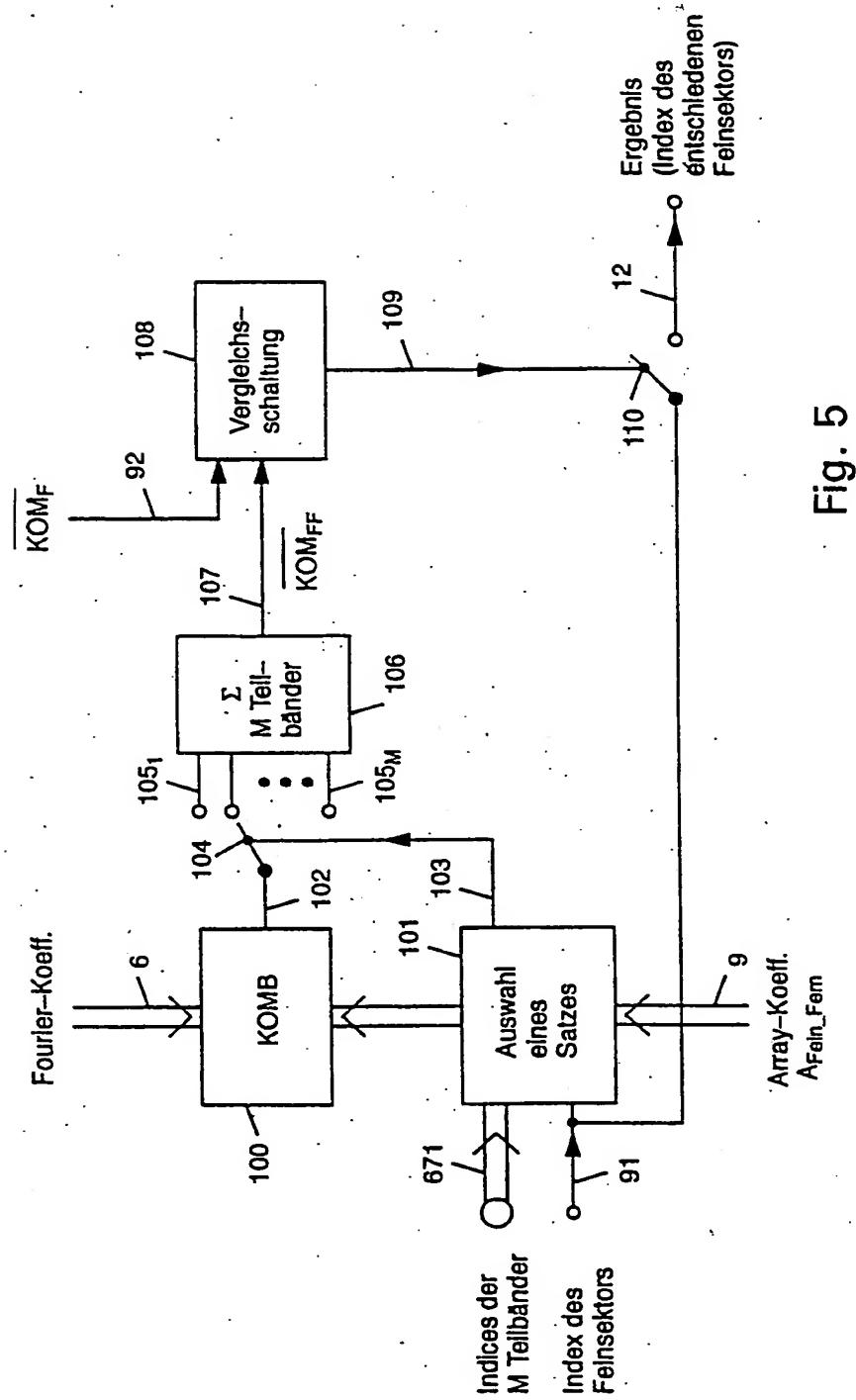


Fig. 5